ELSEVIER
Journal of Computational and Applied Mathematics 116 (2000) 93–104

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

Use of Rényi's divergence to test for the equality of the coefficients of variation ☆

M.C. Pardo *, J.A. Pardo

Faculty of Mathematics, Department of Statistics and O.R., Complutense University of Madrid, Av. Complutense s/n, 28040-Madrid, Spain

Received 9 March 1999; received in revised form 2 August 1999

Abstract

A new family of test statistics based on Rényi's divergence is introduced for the hypothesis that the coefficients of variation of k normal populations are equal. A comparative simulation study is carried out concerning the size and power of these test statistics and earlier ones. Finally, two members of the new family of tests emerge as the best from the simulation study.

Keywords: Coefficient of variation; Rényi's divergence; Simulation; Size; Power

1. Introduction

The coefficient of variation of a random variable provides a unitless measure of relative variability. It is very important in physical, biological, medical and financial sciences. Populations can have the same relative variability even if the means and variances of the variable of interest are different. In situations like this it might be possible to transform the dependent variable so that the variances are similar and then use ANOVA to compare the means. However, there are occasions when it is not possible to find a transformation which will make the assumption of equal variances acceptable or when interest is in a comparison of relative variability. In these situations a test for the equality of coefficients of variation is a reasonable approach.

Miller and Karson [12] presented the likelihood ratio test for the equality of two coefficients of variation. Doornbos and Dijkstra [3] extended this result and developed the so-called noncentral

☆ This work was supported by grant DGICYT PB96-0635.

* Corresponding author. Tel.: +34-913-944473; fax:+34-913-944607.
E-mail address: mcpardo@eucmax.sim.ucm.es (M.C. Pardo)

PII: S0377-0427(99)00312-X

t-test for the equality of the coefficients of variation from k normal populations. This last one was based upon the sample coefficient of variation. Bennett [1] presented a likelihood ratio test which uses an approximation to the distribution of the sample coefficient of variation obtained by Mckay [10] for k normal populations. Shafer and Sullivan [17] presented a modified version of Bennett's test which was motivated by work by Iglewicz and Myers [7]. Miller [11] and Feltz and Miller [4] derived one, two and k-sample tests for coefficients of variation of normal populations based on the fact that the sample coefficient of variation computed from a sample drawn from a normally distributed population is asymptotically normal. Rao and Vidya [15] provided a Wald test for testing the equality of coefficients of variation in two populations with equal sample sizes. More recently, Gupta and Ma [6] developed one new test, the so-called score test and also extended the Wald test to more than two populations and to samples of possibly unequal sizes.

In Section 2, we review several of the above-mentioned tests. A new test statistic based on Rényi's divergence is proposed in Section 3. Finally, results from a simulation study evaluating the sizes and powers of all test statistics of Sections 2 and 3 will be presented in Section 4. This study is carried out under normality assumptions as well as under nonnormality.

2. Background

Let $(X_{i1},\dots,X_{in_i},\ i=1,\dots,k)$ represent k independent normal random samples and assume that $E[X_{ij}]=\mu_i$ and $V[X_{ij}]=\sigma_i^2$ for $i=1,\dots,k$ and $j=1,\dots,n_i$. The coefficient of variation for population i is $R_i=\sigma_i/\mu_i,\ i=1,\dots,k$. As noted by Johnson and Welch [8] and Koopmans et al. [9], in most practical cases where the coefficient of variation is of interest, the random variable is positive, and therefore it shall be assumed that $\mu_i>0$ and hence $R_i>0$.

It is desired to test the null hypothesis

$$\mathrm{H}_0:\ R_i=R,\ i=1,\dots,k;\ R \text{ unknown}$$

against

$$\mathrm{H}_1:\ R_i\neq R_j,\ i\neq j \text{ for at least one pair } (i,j) \text{ where } i,j\in\{1,\dots,k\}.$$

Bennett [1] presented the following test statistic for this problem:

$$B=(n-k)\ln\left(\sum_{i=1}^{k}\frac{n_i(S_{(i)}/\bar{X}_i)^2}{(n-k)(1+(S_{(i)}/\bar{X}_i)^2)}\right)-\sum_{i=1}^{k}(n_i-1)\ln\left(\frac{n_i(S_{(i)}/\bar{X}_i)^2}{(n_i-1)(1+(S_{(i)}/\bar{X}_i)^2)}\right)$$

and the modified Bennett's test proposed by Shafer and Sullivan [17] is as follows:

$$\mathrm{MB}=(n-k)\ln\left(\sum_{i=1}^{k}\frac{n_i(S_i/\bar{X}_i)^2}{(n-k)(1+(S_i/\bar{X}_i)^2)}\right)-\sum_{i=1}^{k}(n_i-1)\ln\left(\frac{n_i(S_i/\bar{X}_i)^2}{(n_i-1)(1+(S_i/\bar{X}_i)^2)}\right),$$

where $n=\sum_{i=1}^{k}n_i,\ S_i^2=\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)^2/n_i,\ S_{(i)}^2=\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)^2/(n_i-1)$ with $i\in\{1,\dots,k\}$. Both statistics are approximately chi-squared distributed with $k-1$ degrees of freedom under the null-hypothesis (see e.g., [18]).

Miller [11] proposed a test for the coefficients of variation of normal populations based on the fact that the coefficient of variation computed from a sample drawn from a normally distributed population is asymptotically normal. Under H_0 and $k=2$, the following asymptotically standard normal test statistic may be used:

$$M_2 = \frac{S_1/\bar{X}_1 - S_2/\bar{X}_2}{\{(1/(n_1-1)+1/(n_2-1))R^2[0.5+R^2]\}^{1/2}}.$$

For k samples, an asymptotically central chi-square with $k-1$ degrees of freedom test statistic is proposed:

$$M_k = \{R^2[0.5+R^2]\}^{-1}\left[\sum_{i=1}^{k}(n_i-1)\left(\frac{S_i}{\bar{X}_i}\right)^2 - \frac{1}{N}\left(\sum_{i=1}^{k}(n_i-1)\frac{S_i}{\bar{X}_i}\right)^2\right].$$

In practice one must estimate R, presumably by

$$R = \frac{1}{n-k}\sum_{i=1}^{k}(n_i-1)\frac{S_i}{\bar{X}_i}.$$

Gupta and Ma [6] generalized the results of Rao and Vidya [15] for $k=4$ and unequal sizes in relation to the Wald test statistic for testing equality of coefficients of variation. This test statistic is given by

$$W = h^{t}(\hat{\theta})[H(\hat{\theta})^{t}I(\hat{\theta})^{-1}H(\hat{\theta})]^{-1}h(\hat{\theta}),$$

where $h(\theta)=(h_1(\theta),\ldots,h_{k-1}(\theta))$ with $h_i(\theta)=(\sigma_i/\mu_i)-(\sigma_{i+1}/\mu_{i+1})$, $H(\theta)=(\partial h_j(\theta)/\partial\theta_i)_{i=1,\ldots,2k;\,j=1,\ldots,k-1}$, $I(\theta)$ the Fisher information matrix and $\hat{\theta}$ is the unrestricted maximum likelihood estimator (MLE) of $\theta=(\mu_1,\sigma_1,\mu_2,\ldots,\mu_k,\sigma_k)$. W is asymptotically distributed as chi-squared with $k-1$ degrees of freedom under the null-hypothesis (see e.g., [18]).

The Score test introduced by Gupta and Ma [6] is given by

$$S = \frac{\tilde{R}^2(2\tilde{R}^2+1)}{2}\sum_{i=1}^{k}\frac{1}{n_i}\left(\frac{\sum_{j=1}^{n_i}(X_{ij}-\tilde{\mu}_i)^2}{\tilde{\mu}_i^2\tilde{R}^3} - \frac{n_i}{\tilde{R}}\right)^2, \tag{1}$$

where $(\tilde{R},\tilde{\mu}_1,\ldots,\tilde{\mu}_k)$ are the restricted maximum likelihood estimators (RMLE) of $(R,\mu_1,\ldots,\mu_k)$ under the null hypothesis. These are obtained by solving the likelihood equations under H_0. Simplifying these equations, we have that

$$\sum_{i=1}^{k}\frac{n_i(1+\sqrt{1+4(1+(S_i/\bar{X}_i)^2)R^2})}{2(1+(S_i/\bar{X}_i)^2)} - n = 0 \tag{2}$$

and

$$\mu_i = \left(\frac{1+\sqrt{1+4(1+(S_i/\bar{X}_i)^2)R^2}}{2(1+(S_i/\bar{X}_i)^2)\bar{X}_i}\right)^{-1}, \quad i=1,\ldots,k. \tag{3}$$

For $k=2$, Gerig and Sen [5] obtained the explicit expressions of the RMLE of $\tilde{R}$, $\tilde{\mu}_1$ and $\tilde{\mu}_2$ but when $k>2$ a numerical method is necessary to get the estimations. The test statistic S, given in 1, is asymptotically distributed as chi-squared with $k-1$ degrees of freedom under the null hypothesis. Then we must reject the null hypothesis at a level α if B, MB, M_k $(k>2)$, W or S are greater than the $1-\alpha$ percentile of the chi-square distribution with $k-1$ degrees of freedom, $\chi^2_{k-1,\alpha}$ and if M_2 is greater than the $1-\alpha$ percentile of the standard normal.

3. Test statistic based on Rényi's divergence

The idea of using divergence measures in testing statistical hypotheses has received a lot of attention in the last years. The divergence statistics obtained by replacing unknown parameters by suitable estimates, have become successful competitors to the classical likelihood ratio-based statistic for testing general composite hypotheses (see, e.g., [14] and further references therein). A divergence is a distance, in a wide sense, between two populations. There are many important families of divergences whose properties have been studied by different authors.

In this paper we consider the Rényi's divergence [16] to define a new family of test statistics for testing equality of coefficients of variation. Recently, this divergence has been used to test equality of variances [13].

Let $(\mathscr{X},\beta_{\mathscr{X}},P_\theta)_{\theta\in\Theta}$ be a measurable space, where $\mathscr{X}\subset R$ is the sample space, $\beta_{\mathscr{X}}$ the corresponding σ-field and $\Theta\subset R^t, t\geqslant 1$. Assume that measures P_θ can be described by densities $f_\theta(x)=(\mathrm{d}P_\theta/\mathrm{d}\mu)(x)$ w.r.t. a dominating σ-finite measure μ on $\mathscr{X}$. The Rényi's divergence for arbitrary densities f_{θ_1} and f_{θ_2} belonging to the family $\{f_\theta, \theta\in\Theta\}$, is given by

$$D_r(\theta_1,\theta_2)=\frac{1}{r(r-1)}\ln\int_{\mathscr{X}} f_{\theta_1}(x)^r f_{\theta_2}(x)^{1-r}\,\mathrm{d}\mu$$

if $r\notin\{0,1\}$, and limiting cases for $r=0$ and 1. That is,

$$D_1(\theta_1,\theta_2)=\lim_{r\uparrow 1} D_r(\theta_1,\theta_2)=\int_{\mathscr{X}} f_{\theta_1}(x)\ln\frac{f_{\theta_1}(x)}{f_{\theta_2}(x)}\,\mathrm{d}\mu,$$

$$D_0(\theta_1,\theta_2)=\lim_{r\downarrow 0} D_r(\theta_1,\theta_2)=\int_{\mathscr{X}} f_{\theta_2}(x)\ln\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)}\,\mathrm{d}\mu=D_1(\theta_2,\theta_1).$$

The measures of divergences $D_1(\theta_1,\theta_2)$ and $D_0(\theta_1,\theta_2)$ are called Kullback–Leibler divergence and reversed Kullback–Leibler divergence, respectively.

Morales et al. [14] studied the problem of testing composite hypothesis H_0: $\theta\in\Theta_0\subset\Theta$ versus H_1: $\theta\in\Theta-\Theta_0$ on the basis of Rényi's divergence using the statistic

$$S_n^r=2nD_r(\hat{\theta}_n,\tilde{\theta}_n),$$

where $\hat{\theta}_n$ is the MLE of $\theta\in\Theta$ and $\tilde{\theta}_n$ is the RMLE with values limited to the hypothesis subset Θ_0. Under standard regularity assumptions, they established that S_n^r is asymptotically distributed chi-squared with d_0 degrees of freedom, where d_0 is the difference of dimensions between Θ and Θ_0. For large n, when $S_n^r=t$, H_0 should be rejected at a level α if $P(\chi^2_{d_0}>t)\leqslant\alpha$ where by $\chi^2_{d_0}$ we denote the chi-squared random variable with d_0 degrees of freedom.

Assume that we are interested in testing a composite hypothesis H_0 about parameters from k populations with distributional structure differing only by means of a parameter θ. Let us denote $\theta_1,\dots,\theta_k$ to the parameter values at populations $1,\dots,k$, respectively. If samples of sizes $n_1,\dots,n_k$ are taken at random and the MLE $\hat{\theta}_1,\dots,\hat{\theta}_k$ and the RMLE $\tilde{\theta}_1,\dots,\tilde{\theta}_k$ are computed, then testing procedure given in [14] is applicable to the case of balanced problems (i.e. problems with $n_1=\cdots=n_k$). Statistic S_n^r is obtained by calculating Rényi's divergence between joint densities

$$\prod_{i=1}^{k} f_{\hat{\theta}_i}(y_i) \quad \text{and} \quad \prod_{i=1}^{k} f_{\tilde{\theta}_i}(y_i).$$

When dealing with k samples of different sizes, S_n^r cannot be used unless it was generalized in some sense. Morales et al. [13] use Rényi's divergence between the estimated likelihoods

$$\prod_{i=1}^{k}\prod_{j=1}^{n_i} f_{\hat{\theta}_i}(x_{ij}) \quad \text{and} \quad \prod_{i=1}^{k}\prod_{j=1}^{n_i} f_{\tilde{\theta}_i}(x_{ij})$$

to define a new test statistic for the case of problems with several populations and unequal sample sizes.

Let $(\mathscr{X}_1,\mathscr{B}_{\mathscr{X}_1},P_{\theta_1})_{\theta_1\in\Theta},\dots,(\mathscr{X}_k,\mathscr{B}_{\mathscr{X}_k},P_{\theta_k})_{\theta_k\in\Theta}$ be k statistical spaces associated to independent populations, where $\mathscr{X}_1=\cdots=\mathscr{X}_k=\mathscr{X}\subset R^k$ are the sample spaces, $\mathscr{B}_{\mathscr{X}_1}=\cdots=\mathscr{B}_{\mathscr{X}_k}=\mathscr{B}_{\mathscr{X}}$ the corresponding Borel σ-fields and $\Theta\subset R^d$ is open. Measure P_{θ_i} is assumed to be described by density $f_{\theta_i}(x)=(\mathrm{d}P_{\theta_i}/\mathrm{d}\mu)(x)$, $i=1,\dots,k$, w.r.t. a dominating measure μ on $\mathscr{X}$. We are interested in testing composite hypotheses concerning the above k populations and based on k independent samples $\boldsymbol{X}_1^{(n_1)}=(X_{11},\dots,X_{1n_1}),\dots,\boldsymbol{X}_k^{(n_k)}=(X_{k1},\dots,X_{kn_k})$. Let us write $\Gamma=\Theta^k\subset R^{dk}$, $\gamma=(\theta_1,\dots,\theta_k)\in\Gamma$, $\theta_i=(\theta_{i,1},\dots,\theta_{i,d})$, $(\mathscr{X}_1\times\cdots\times\mathscr{X}_1,\sigma(\mathscr{B}_{\mathscr{X}_1}\times\cdots\times\mathscr{B}_{\mathscr{X}_k}),P_{\theta_1}\otimes\cdots\otimes P_{\theta_k})_{(\theta_1,\dots,\theta_k)\in\Gamma}$, for the product statistical space and $f_\gamma(x_1,\dots,x_k)=\prod_{i=1}^k f_{\theta_i}(x_i)$, for the density of $P_{\theta_1}\otimes\cdots\otimes P_{\theta_k}$ w.r.t. the product measure μ^k. Let $(\mathscr{Z},\mathscr{B}_{\mathscr{Z}},P_\gamma^{(\boldsymbol{n})})$ be a statistical space where $\mathscr{Z}=\mathscr{X}_1^{n_1}\times\cdots\times\mathscr{X}_k^{n_k}$, $\mathscr{B}_{\mathscr{Z}}=\sigma(\mathscr{B}_{\mathscr{X}_1^{n_1}}\times\cdots\times\mathscr{B}_{\mathscr{X}_k^{n_k}})$ and $P_\gamma^{(\boldsymbol{n})}=P_{\theta_1}^{n_1}\otimes\cdots\otimes P_{\theta_k}^{n_k}$. Let us consider the random vector $\boldsymbol{Z}^{(\boldsymbol{n})}=(\boldsymbol{X}_1^{(n_1)},\dots,\boldsymbol{X}_k^{(n_k)})$ with realizations $z=(x_{11},\dots,x_{1n_1},\dots,x_{k1},\dots,x_{kn_k})$. The statistic

$$T_{\boldsymbol{n}}^r=2D_r(\hat{\gamma}_{\boldsymbol{n}},\tilde{\gamma}_{\boldsymbol{n}})=\frac{2}{r(r-1)}\ln\int_{\mathscr{X}^n} g_{\hat{\gamma}_{\boldsymbol{n}}}(z)^r g_{\tilde{\gamma}_{\boldsymbol{n}}}(z)^{1-r}\,\mathrm{d}\mu^n$$

and limiting cases for $r=0,1$ was proposed by Morales et al. [13] for testing the hypothesis $\mathrm{H}_0\equiv\Gamma_0\subset\Gamma$ with $\Gamma=\Theta^k\subset R^k$ where $\boldsymbol{n}=(n_1,\dots,n_k)$, $\hat{\gamma}_{\boldsymbol{n}}\equiv$ MLE of γ in Γ and $\tilde{\gamma}_{\boldsymbol{n}}\equiv$ MLE of γ in Γ_0. Under standard regularity assumptions they established that $T_{\boldsymbol{n}}^r$ is asymptotically chi-squared distributed with d_0 degrees of freedom, where d_0 is the difference between the dimension of the parameter space Γ and the hypothesis space Γ_0. In the particular but important case of exponential family models, the test statistic $T_{\boldsymbol{n}}^1=\lim_{r\uparrow 1}T_{\boldsymbol{n}}^r$ for testing any hypothesis coincides with the likelihood ratio test statistic when the exponential family is not overparametrized.

In our case, we are interested in testing

$$\mathrm{H}_0\colon R_1=\cdots=R_k=R \quad (R \text{ unknown}).$$

Let $X_{i1},\dots,X_{in_i}$, $i=1,\dots,k$ be k independent samples from normal distributions. The joint parameter space is

$$\Gamma=\{(x_1,\dots,x_k,y_1,\dots,y_k)/x_i\in R, y_i\in R \text{ and } y_i>0, i=1,\dots,k\}$$

and its restriction to H_0 is

$$\Gamma_0 = \left\{ (x_1,\ldots,x_k,y_1,\ldots,y_k) \in \Gamma \Bigg/ \frac{x_1}{\sqrt{y_1}} = \cdots = \frac{x_k}{\sqrt{y_k}} \right\}.$$

So we introduce the new family of test statistics for testing the equality of k coefficients of variation in k normal populations as

$$T_{\boldsymbol{n}}^r = 2D_r((\hat{\mu},\hat{\Sigma}),(\tilde{\mu},\tilde{\Sigma})),$$

where

$$\hat{\mu} = (\bar{X}_1, ..^{n_1)} .., \bar{X}_1, \ldots, \bar{X}_k, ..^{n_k)} .. \bar{X}_k) \quad \text{and} \quad \hat{\Sigma} = \operatorname{diag}(S_1^2, ..^{n_1)} \ldots, S_1^2, \ldots, S_k^2, ..^{n_k)} \ldots, S_k^2)$$

are the MLE of μ and Σ and

$$\tilde{\mu} = (\tilde{\mu}_1, ..^{n_1)} .., \tilde{\mu}_1, \ldots, \tilde{\mu}_k, ..^{n_k)} .., \tilde{\mu}_k) \quad \text{and} \quad \tilde{\Sigma} = \operatorname{diag}(\tilde{\mu}_1^2\tilde{R}^2, ..^{n_1)} .., \tilde{\mu}_1^2\tilde{R}^2, \ldots, \tilde{\mu}_k^2\tilde{R}^2, ..^{n_k)} .., \tilde{\mu}_k^2\tilde{R}^2)$$

the RMLE of μ and Σ. This test statistic is asymptotically distributed chi-squared with $k-1$ degrees of freedom under the null hypothesis since $d_0 = k-1$. Therefore, an asymptotically α-level test for the problem of testing the equality of coefficients of variation would reject H_0 if $T_{\boldsymbol{n}}^r > \chi^2_{k-1,\alpha}$.

Using Rényi's divergence, given in [2], between two k-dimensional normal populations we obtain the following expression for $T_{\boldsymbol{n}}^r$:

$$T_{\boldsymbol{n}}^r = \sum_{i=1}^k n_i \left(\frac{(\bar{X}_i - \tilde{\mu}_i)^2}{(1-r)S_i^2 + r\tilde{\mu}_i^2\tilde{R}^2} + \frac{1}{r(1-r)} \ln \frac{(1-r)S_i^2 + r\tilde{\mu}_i^2\tilde{R}^2}{(S_i^2)^{1-r} + (\tilde{\mu}_i^2\tilde{R}^2)^r} \right)$$

if $r \notin \{0,1\}$, and limiting cases for $r=0$ and 1 are given by

$$T_{\boldsymbol{n}}^1 = \lim_{r\uparrow 1} T_{\boldsymbol{n}}^r = \sum_{i=1}^k n_i \ln \frac{\tilde{\mu}_i^2\tilde{R}^2}{S_i^2},$$

$$T_{\boldsymbol{n}}^0 = \lim_{r\downarrow 0} T_{\boldsymbol{n}}^r = \sum_{i=1}^k n_i \left(\frac{(\bar{X}_i - \tilde{\mu}_i)^2}{S_i^2} + \frac{\tilde{\mu}_i^2\tilde{R}^2}{S_i^2} - 1 + \ln \frac{S_i^2}{\tilde{\mu}_i^2\tilde{R}^2} \right).$$

By solving Eqs. (2) and (3), we get $(\tilde{R}, \tilde{\mu}_1, \ldots, \tilde{\mu}_k)$. Since Eq. (2) cannot be solved algebraically when $k > 2$, we use a numerical method given by Gupta and Ma [6] to solve it.

Note that $T_{\boldsymbol{n}}^1$ coincides with the likelihood ratio test since the normal distribution belongs to a not overparametrized exponential family. Firstly, it was studied by Miller and Karson [12] for testing equality of coefficients of variation for $k=2$ and equal sample sizes. Later, it was generalized by Doornbos and Dijkstra [3] for $k>2$ and unequal sample sizes. Now it emerges as a particular case of our new family of test statistics. This allows to study it jointly as a member of the family of test statistics $T_{\boldsymbol{n}}^r$.

4. Simulation results

Monte Carlo experiments were performed to evaluate several members of the family of test statistics $T_{\boldsymbol{n}}^r$ and the Bennett, modified Bennett, Miller, Wald and Score test statistics in terms of

Table 1
Estimated size for four normal populations

	$n_1 = n_2 = n_3 = n_4 = 10$						$n_1 = 5, n_2 = 10, n_3 = 15, n_4 = 25$					
R	0.1	1/3	0.5	1	1.5	2	0.1	1/3	0.5	1	1.5	2
B	0.055	0.050	0.039	0.022	0.004	0.000	0.061	0.046	0.039	0.016	0.004	0.002
MB	0.056	0.050	0.041	0.024	0.006	0.000	0.059	0.047	0.043	0.021	0.009	0.004
M	0.053	0.046	0.042	0.042	0.006	0.002	0.049	0.039	0.052	0.085	0.059	0.018
W	0.084	0.070	0.054	0.014	0.000	0.000	0.173	0.156	0.149	0.089	0.037	0.013
S	0.048	0.051	0.039	0.054	0.031	0.012	0.043	0.036	0.047	0.040	0.023	0.017
T_n^{-1}	0.425	0.420	0.404	0.378	0.311	0.250	0.486	0.472	0.448	0.407	0.338	0.313
$T_n^{-0.3}$	0.316	0.312	0.315	0.293	0.244	0.187	0.370	0.357	0.337	0.308	0.280	0.252
T_n^{0}	0.182	0.186	0.172	0.181	0.152	0.117	0.222	0.215	0.210	0.198	0.178	0.162
$T_n^{0.5}$	0.111	0.106	0.101	0.114	0.099	0.073	0.130	0.127	0.138	0.131	0.112	0.094
T_n^{1}	0.071	0.068	0.065	0.079	0.064	0.042	0.083	0.076	0.085	0.090	0.073	0.044
$T_n^{1.3}$	0.069	0.064	0.062	0.072	0.051	0.031	0.076	0.073	0.075	0.063	0.058	0.033
$T_n^{1.6}$	0.081	0.080	0.072	0.067	0.043	0.020	0.091	0.075	0.075	0.053	0.041	0.024
T_n^{2}	0.143	0.134	0.104	0.068	0.030	0.014	0.139	0.115	0.102	0.052	0.032	0.020

size and power. Random samples were generated for four populations ($k = 4$) using the normal, gamma, log-normal and uniform distributions. Each Monte Carlo experiment consisted of 1000 replications. Two sample size patterns were used for the populations. We count the number of times for each test that the null hypothesis or alternative hypothesis were accepted, to obtain the size or the power, respectively. The critical values are the $1 - \alpha$ percentile of the chi-square distribution with $k - 1$ degrees of freedom for all test statistics considered.

Table 1 presents the simulation results corresponding to estimated size for the statistics for four normal populations for H_0: $R_i = R$, $i = 1, 2, 3, 4$, for six values of R. The nominal size was set at 0.05. It can be seen from this table that the estimated size using the Bennett, modified Bennett and Miller tests is close to 0.05 for both sample sizes when $0 < R < 1$. However, these tests would be inappropriate for R values greater or equal than one since they are much too conservative. Although $T_n^{1.3}$ and $T_n^{1.6}$ are worse than these tests for $R < 1$ because they are more liberal. However, they work well for $R \geqslant 1$ and not bad for the rest. These tests become a slight conservative when $R = 2$ but not quite as conservative as the B, MB and M tests.

Tables 2–4 present the results of the estimated size for the gamma, log-normal and uniform distributions for situations similar to those used for the normal.

The conclusions for the gamma and log-normal are analogous from those for the normal when $R < 1$. The poor results for the uniform imply that most of the tests are inappropriate for uniform populations. However, the estimated size using T_n^{1} and $T_n^{1.3}$ test statistics are close to 0.05 for both sample sizes when $R \geqslant 1$ and T_n^{0} and $T_n^{0.5}$ perform well when $R < 1$ for equal and unequal sizes, respectively.

Tables 5–8 include the estimated power for four populations given the following alternative hypotheses:

Table 2
Estimated size for four gamma populations

	$n_1=n_2=n_3=n_4=10$						$n_1=5,\ n_2=10,\ n_3=15,\ n_4=25$					
R	0.1	1/3	0.5	1	1.5	2	0.1	1/3	0.5	1	1.5	2
B	0.048	0.031	0.028	0.004	0.000	0.000	0.048	0.036	0.024	0.003	0.000	0.000
MB	0.048	0.032	0.030	0.005	0.000	0.000	0.050	0.036	0.031	0.005	0.000	0.001
M	0.047	0.033	0.024	0.003	0.000	0.000	0.039	0.033	0.017	0.005	0.000	0.000
W	0.082	0.039	0.035	0.009	0.000	0.000	0.192	0.134	0.115	0.059	0.000	0.013
S	0.050	0.037	0.024	0.004	0.000	0.000	0.035	0.035	0.013	0.005	0.000	0.000
T_n^{-1}	0.424	0.349	0.325	0.197	0.009	0.031	0.496	0.452	0.366	0.270	0.050	0.095
$T_n^{-0.3}$	0.318	0.241	0.224	0.116	0.003	0.015	0.382	0.335	0.280	0.180	0.019	0.063
T_n^{0}	0.177	0.119	0.114	0.051	0.001	0.004	0.244	0.186	0.149	0.086	0.003	0.020
$T_n^{0.5}$	0.108	0.065	0.065	0.020	0.000	0.001	0.148	0.108	0.085	0.042	0.001	0.009
T_n^{1}	0.071	0.048	0.040	0.010	0.000	0.000	0.097	0.062	0.049	0.021	0.000	0.003
$T_n^{1.3}$	0.076	0.043	0.039	0.010	0.000	0.000	0.080	0.053	0.043	0.014	0.000	0.001
$T_n^{1.6}$	0.085	0.051	0.046	0.010	0.000	0.000	0.082	0.058	0.041	0.011	0.000	0.001
T_n^{2}	0.124	0.089	0.071	0.017	0.000	0.000	0.127	0.086	0.050	0.012	0.000	0.001

Table 3
Estimated size for four lognormal populations

	$n_1=n_2=n_3=n_4=10$						$n_1=5,\ n_2=10,\ n_3=15,\ n_4=25$					
R	0.1	1/3	0.5	1	1.5	2	0.1	1/3	0.5	1	1.5	2
B	0.060	0.051	0.029	0.000	0.002	0.000	0.058	0.046	0.032	0.010	0.002	0.000
MB	0.060	0.054	0.034	0.000	0.003	0.001	0.058	0.051	0.050	0.016	0.007	0.002
M	0.057	0.056	0.041	0.007	0.006	0.000	0.049	0.047	0.036	0.012	0.005	0.001
W	0.082	0.056	0.036	0.002	0.001	0.000	0.175	0.138	0.130	0.090	0.036	0.032
S	0.048	0.058	0.041	0.007	0.006	0.000	0.041	0.049	0.038	0.015	0.005	0.004
T_n^{-1}	0.433	0.401	0.361	0.247	0.158	0.120	0.480	0.475	0.444	0.394	0.331	0.299
$T_n^{-0.3}$	0.319	0.290	0.257	0.163	0.097	0.066	0.376	0.377	0.347	0.308	0.239	0.218
T_n^{0}	0.182	0.164	0.137	0.084	0.040	0.028	0.219	0.191	0.204	0.165	0.113	0.113
$T_n^{0.5}$	0.099	0.098	0.082	0.036	0.021	0.010	0.129	0.121	0.134	0.097	0.060	0.055
T_n^{1}	0.071	0.078	0.054	0.019	0.014	0.006	0.084	0.084	0.091	0.057	0.029	0.031
$T_n^{1.3}$	0.073	0.074	0.051	0.013	0.010	0.004	0.072	0.077	0.075	0.037	0.023	0.021
$T_n^{1.6}$	0.083	0.084	0.064	0.012	0.009	0.003	0.086	0.088	0.081	0.029	0.021	0.021
T_n^{2}	0.142	0.125	0.097	0.019	0.009	0.003	0.137	0.115	0.091	0.027	0.016	0.013

$$\mathrm{H}_1':\quad R_1=R_2=0.1;\ R_3=R_4=1,$$

$$\mathrm{H}_2':\quad R_1=R_2=R_3=0.1;\ R_4=1,$$

$$\mathrm{H}_3':\quad R_1=0.1;\ R_2=1/3;\ R_3=0.5;\ R_4=1,$$

$$\mathrm{H}_4':\quad R_1=0.5;\ R_2=R_3=0.1;\ R_4=\tfrac{1}{3},$$

$$\mathrm{H}_5':\quad R_1=R_2=0.5;\ R_3=R_4=\tfrac{1}{3}.$$

Table 4
Estimated size for four uniform populations

	$n_1 = n_2 = n_3 = n_4 = 10$						$n_1 = 5,\ n_2 = 10,\ n_3 = 15,\ n_4 = 25$					
R	0.1	1/3	0.5	1	1.5	2	0.1	1/3	0.5	1	1.5	2
B	0.007	0.013	0.014	0.011	0.010	0.007	0.006	0.008	0.014	0.011	0.008	0.000
MB	0.007	0.014	0.017	0.013	0.012	0.008	0.006	0.015	0.018	0.018	0.010	0.000
M	0.004	0.009	0.004	0.018	0.007	0.002	0.004	0.004	0.007	0.060	0.061	0.034
W	0.032	0.029	0.026	0.012	0.000	0.000	0.085	0.093	0.116	0.068	0.037	0.010
S	0.005	0.008	0.004	0.021	0.029	0.016	0.001	0.002	0.005	0.023	0.012	0.010
T_n^{-1}	0.211	0.199	0.255	0.289	0.250	0.225	0.328	0.303	0.302	0.281	0.287	0.246
$T_n^{-0.3}$	0.123	0.111	0.162	0.201	0.199	0.173	0.220	0.216	0.218	0.217	0.227	0.186
T_n^{0}	0.056	0.066	0.079	0.129	0.123	0.103	0.108	0.108	0.138	0.131	0.144	0.120
$T_n^{0.5}$	0.021	0.037	0.036	0.082	0.082	0.069	0.039	0.049	0.070	0.083	0.085	0.073
T_n^{1}	0.013	0.022	0.025	0.053	0.059	0.046	0.014	0.024	0.035	0.056	0.050	0.044
$T_n^{1.3}$	0.010	0.019	0.021	0.043	0.047	0.037	0.008	0.020	0.028	0.045	0.042	0.031
$T_n^{1.6}$	0.011	0.020	0.019	0.034	0.038	0.031	0.007	0.012	0.023	0.038	0.028	0.025
T_n^{2}	0.019	0.029	0.028	0.038	0.033	0.025	0.024	0.029	0.026	0.034	0.020	0.014

Table 5
Estimated power for four normal populations

	$n_1 = n_2 = n_3 = n_4 = 10$					$n_1 = 5,\ n_2 = 10,\ n_3 = 15,\ n_4 = 25$				
	H_1'	H_2'	H_3'	H_4'	H_5'	H_1'	H_2'	H_3'	H_4'	H_5'
B	1.000	1.000	0.998	0.998	0.202	1.000	1.000	0.996	0.318	0.318
MB	1.000	1.000	0.998	0.998	0.215	1.000	1.000	0.998	1.000	0.284
M	0.997	0.997	0.996	0.998	0.212	0.989	1.000	0.903	1.000	0.368
W	0.924	0.524	1.000	0.995	0.151	1.000	0.984	1.000	1.000	0.078
S	1.000	1.000	0.966	0.998	0.205	1.000	1.000	0.881	1.000	0.367
T_n^{-1}	1.000	1.000	1.000	1.000	0.660	1.000	1.000	1.000	1.000	0.502
$T_n^{-0.3}$	1.000	1.000	1.000	1.000	0.565	1.000	1.000	1.000	1.000	0.422
T_n^{0}	1.000	1.000	1.000	1.000	0.424	1.000	1.000	1.000	1.000	0.312
$T_n^{0.5}$	1.000	1.000	1.000	1.000	0.321	1.000	1.000	1.000	1.000	0.276
T_n^{1}	1.000	1.000	0.999	1.000	0.267	1.000	1.000	1.000	1.000	0.290
$T_n^{1.3}$	1.000	1.000	0.998	1.000	0.256	1.000	1.000	1.000	1.000	0.330
$T_n^{1.6}$	1.000	1.000	0.998	1.000	0.271	1.000	1.000	0.999	1.000	0.413
T_n^{2}	1.000	1.000	0.997	1.000	0.352	1.000	1.000	0.998	1.000	0.522

The two Rényi statistics, $T_n^{1.3}$ and $T_n^{1.6}$, whose estimated sizes appear very close to the nominal size for normal, gamma and log-normal populations have higher power than B, MB and M tests. If we do not mind use slightly liberal tests in favour of more powerful tests we would use $T_n^{1.3}$ as $T_n^{1.6}$ for all the cases.

The estimated power appears to be very good for each distribution. As expected, the power is largest in all tables for those alternatives which represent more separation of the R_i's and is smallest when the R_i's are close.

Table 6
Estimated power for four gamma populations

	$n_1=n_2=n_3=n_4=10$					$n_1=5,\ n_2=10,\ n_3=15,\ n_4=25$				
	H_1'	H_2'	H_3'	H_4'	H_5'	H_1'	H_2'	H_3'	H_4'	H_5'
B	1.000	1.000	1.000	1.000	0.043	1.000	1.000	0.936	1.000	0.075
MB	1.000	1.000	1.000	1.000	0.047	1.000	1.000	0.971	1.000	0.055
M	1.000	1.000	0.995	1.000	0.056	1.000	1.000	0.674	1.000	0.082
W	0.076	0.883	1.000	0.997	0.048	1.000	1.000	1.000	1.000	0.068
S	1.000	1.000	0.880	0.998	0.058	1.000	1.000	0.573	1.000	0.087
T_n^{-1}	1.000	1.000	1.000	1.000	0.462	1.000	1.000	1.000	1.000	0.319
$T_n^{-0.3}$	1.000	1.000	1.000	1.000	0.329	1.000	1.000	1.000	1.000	0.214
T_n^{0}	1.000	1.000	1.000	1.000	0.181	1.000	1.000	1.000	1.000	0.130
$T_n^{0.5}$	1.000	1.000	1.000	1.000	0.110	1.000	1.000	1.000	1.000	0.090
T_n^{1}	1.000	1.000	1.000	1.000	0.076	1.000	1.000	0.998	1.000	0.073
$T_n^{1.3}$	1.000	1.000	0.999	1.000	0.070	1.000	1.000	0.991	1.000	0.080
$T_n^{1.6}$	1.000	1.000	0.999	1.000	0.080	1.000	1.000	0.980	1.000	0.125
T_n^{2}	1.000	1.000	0.999	1.000	0.133	1.000	1.000	0.953	1.000	0.224

Table 7
Estimated power for four lognormal populations

	$n_1=n_2=n_3=n_4=10$					$n_1=5,\ n_2=10,\ n_3=15,\ n_4=25$				
	H_1'	H_2'	H_3'	H_4'	H_5'	H_1'	H_2'	H_3'	H_4'	H_5'
B	1.000	1.000	0.998	0.998	0.147	1.000	1.000	0.989	1.000	0.208
MB	1.000	1.000	0.998	0.998	0.158	1.000	1.000	0.995	1.000	0.178
M	1.000	1.000	0.998	0.998	0.155	1.000	1.000	0.883	1.000	0.240
W	0.998	0.708	1.000	0.996	0.111	1.000	0.992	1.000	1.000	0.072
S	1.000	1.000	0.946	0.997	0.149	1.000	1.000	0.829	0.999	0.243
T_n^{-1}	1.000	1.000	1.000	1.000	0.607	1.000	1.000	1.000	1.000	0.435
$T_n^{-0.3}$	1.000	1.000	1.000	1.000	0.520	1.000	1.000	1.000	1.000	0.347
T_n^{0}	1.000	1.000	1.000	1.000	0.350	1.000	1.000	1.000	1.000	0.242
$T_n^{0.5}$	1.000	1.000	1.000	1.000	0.251	1.000	1.000	1.000	1.000	0.189
T_n^{1}	1.000	1.000	1.000	1.000	0.207	1.000	1.000	1.000	1.000	0.187
$T_n^{1.3}$	1.000	1.000	0.998	0.999	0.202	1.000	1.000	0.999	1.000	0.223
$T_n^{1.6}$	1.000	1.000	0.999	0.999	0.208	1.000	1.000	0.997	1.000	0.293
T_n^{2}	1.000	1.000	0.999	1.000	0.286	1.000	1.000	0.993	1.000	0.412

Note that the Wald statistic has an unusual low power. Although not reported, simulation runs were performed for more sample sizes. The conclusion is that very small powers are obtained for Wald, Score and Miller tests for some small sample sizes.

Table 8
Estimated power for four uniform populations

	$n_1=n_2=n_3=n_4=10$					$n_1=5,\ n_2=10,\ n_3=15,\ n_4=25$				
	H_1'	H_2'	H_3'	H_4'	H_5'	H_1'	H_2'	H_3'	H_4'	H_5'
B	1.000	1.000	1.000	0.999	0.082	1.000	1.000	0.998	1.000	0.214
MB	1.000	1.000	1.000	0.999	0.089	1.000	1.000	0.999	1.000	0.167
M	0.996	1.000	0.999	0.999	0.092	0.994	1.000	0.922	1.000	0.264
W	0.909	0.358	1.000	0.999	0.069	0.998	0.986	1.000	1.000	0.046
S	1.000	1.000	0.986	0.999	0.094	1.000	1.000	0.899	1.000	0.255
T_n^{-1}	1.000	1.000	1.000	1.000	0.551	1.000	1.000	1.000	1.000	0.343
$T_n^{-0.3}$	1.000	1.000	1.000	1.000	0.424	1.000	1.000	1.000	1.000	0.272
T_n^{0}	1.000	1.000	1.000	1.000	0.253	1.000	1.000	1.000	1.000	0.183
$T_n^{0.5}$	1.000	1.000	1.000	1.000	0.175	1.000	1.000	1.000	1.000	0.160
T_n^{1}	1.000	1.000	1.000	1.000	0.128	1.000	1.000	1.000	1.000	0.166
$T_n^{1.3}$	1.000	1.000	1.000	1.000	0.125	1.000	1.000	0.999	1.000	0.206
$T_n^{1.6}$	1.000	1.000	1.000	1.000	0.144	1.000	1.000	0.999	1.000	0.302
T_n^{2}	1.000	1.000	1.000	1.000	0.210	1.000	1.000	0.998	1.000	0.460

5. Conclusions

A new family of test statistics based on Rényi's divergence, T_n^r, for testing the equality of the coefficients of variation from k normal populations has been introduced and studied in Section 3. Section 4 presents a simulation study of some values of the parameter r ($r=-1,-0.3,0,0.5,1,1.3,1.6,2$) associated to the new family introduced as well as a comparison with the most well-known test statistics for equality of coefficients of variation introduced until now in the literature and developed in Section 2. After the simulation study and for normal populations, we recommend to use $T_n^{1.3}$ or $T_n^{1.6}$ for testing equality of coefficients of variation if $R \geqslant 1$. If $0<R<1$ and we do not mind to lose power in favour of the accuracy of type-I error then the B and MB tests are recommended. Finally, we do not recommend to use the Wald, Score or Miller tests when the sample sizes tend to be smaller since their powers decrease sharply. In addition, we have studied the robustness of these tests under departures from normality. The classic test statistics considered are severely affected when the true distribution is uniform but T_n^1 and $T_n^{1.3}$ appear to be very good when $R \geqslant 1$ and when $R<1$, T_n^0 and $T_n^{0.5}$ emerge as the best for equal and unequal sizes respectively. Finally, $T_n^{1.3}$ or $T_n^{1.6}$ are the best as long as the distribution is approximately bell-shaped when $R<1$ since in other case they are too conservative.

Acknowledgements

We wish to thank to the referee for his very helpful suggestions since they have improved the paper.

References

[1] B.M. Bennett, On an approximate test for homogeneity of coefficients of variation, in: Walter John Ziegler, Contributing to Applied Statistics, Birhauser Verlag, Basel and Stuttgart, 1976, pp. 169–171.

[2] J. Burbea, The convexity with respect to Gaussian distributions of divergences of order α, Utilitas Math. 24 (1982) 171–192.

[3] R. Doornbos, J.B. Dijkstra, A multi sample tests for the equality of coefficients of variation in normal populations, Comm. Statist. Simulation Comput. 12 (2) (1983) 147–158.

[4] C.J. Feltz, G.E. Miller, An asymptotic tests for the equality of coefficients of variation in k populations, Statist. Med. 15 (1996) 647–658.

[5] T.M. Gerig, A.R. Sen, MLE in two normal samples with equal but unknown population coefficients of variation, Comm. Statist. Simulation Comput. 12 (2) (1980) 147–158.

[6] R.C. Gupta, S. Ma, Testing the equality of coefficients of variation in k normal populations, Comm. Statist. Theory Methods 25 (1) (1996) 115–132.

[7] B. Iglewicz, R.H. Myers, Comparisons of approximations to the percentage points of the sample coefficient of variation, Technometrics 12 (1) (1970) 166–169.

[8] N.L. Johnson, B. Welch, Application of the noncentral t distribution, Biometrika 31 (1940) 362–389.

[9] L. Koopmans, D. Owen, J. Rosenblatt, Confidence intervals for the coefficient of variation for normal and lognormal distributions, Biometrika 51 (1964) 25–32.

[10] A.T. McKay, Distribution of the coefficient of variation and the extended t-distribution, J. Roy. Statist. Soc. 95 (1932) 695–698.

[11] G.E. Miller, Asymptotic tests statistics for coefficients of variation, Comm. Statist. Simulation Comput. 20 (10) (1991) 3351–3363.

[12] E.G. Miller, M.J. Karson, Testing equality of two coefficients of variation, American Statistical Association: Proceedings of the Business and Economic Statistics Section, Part. I, 1977, pp. 278–283.

[13] D. Morales, L. Pardo, M.C. Pardo, The behaviour of the Renyi's divergence for homogeneity of variances, Technical Report, Department of Statistics and O.R., Universidad Complutense de Madrid, 1998.

[14] D. Morales, L. Pardo, I. Vajda, Some new statistics for testing hypotheses in parametric models, J. Multivariate Anal. 62 (1997) 137–168.

[15] K.A. Rao, R. Vidya, On the performance of a test for coefficient of variation, Calcutta Statist. Assoc. Bull. 42 (165–166) (1992) 87–95.

[16] A. Rényi, On measures of entropy and information, Proceedings of the Forth Berkeley Symposium on Mathematical Statistics and Probability, 1, 1961, pp. 547–561.

[17] N.J. Shafer, J.A. Sullivan, A simulation study of a test for the equality of the coefficients of variation, Comm. Statist. Simulation Comput. 15 (3) (1986) 681–695.

[18] S.D. Silvey, Statistical Inference, Chapman & Hall, London, 1970.